

# IMPROVISATION AND GESTURE AS FORM DETERMINANTS IN WORKS WITH ELECTRONICS

Alyssa Aska

University of Calgary

alyssa.aska@ucalgary.ca

## ABSTRACT

This paper examines several examples that use electronics as form determinants in works with some degree of structured improvisation. Three works created by the author are discussed, each of which uses gestural controller input to realize an indeterminate form in some way. The application of such principles and systems to venues such as networked performance is explored. While each of these works contains an improvisatory and/or aleatoric element, much of their content is composed, which brings the role of the composer into question. The “improviser”, who in these works advances the work temporally and determines the overall form, is actually taking on the more familiar role of the conductor. Therefore, these works also bring up important conversation topics regarding performance practice in works that contain electronics and how they are realized.

## 1. INTRODUCTION

Improvisation in some form permeates nearly all genres and eras of music: the baroque had improvised keyboard music, improvisatory solos are an integral part of jazz realization, and even classical concerti contain cadenza sections, which are improvised by a soloist. When one considers the ornamentations that were assumed in certain styles of baroque music, as well as temporal improvisation such as rubato in romantic piano music, it is apparent that some degree of performance liberty, or interpretation, is applied to nearly all music, no matter how precisely the notation of the work. Electronic systems, particularly live ones, often have some elements that change, even if slightly, based on performance conditions. This is true of acoustic instruments to an extent (elements such as reverb and delay are dependent on venue, and musicians may compensate for such and perform differently) as well. Taking into account computational processing lag, minute differences in algorithm realization, and live interactive components that may differ in each performance, it is clear that the use of electronics is adapted well to flexible performance realization. It is this flexibility requirement that led composers to seek live systems in favour of fixed media coupled with instruments in the past – for many works that included tape alongside live instruments, there simply wasn’t enough room for the performer to give a fluid performance whilst also adhering to the strict time of the tape. (This is not the case,

however, for all tape music, when one considers some built-in flexibility in works such as Mario Davidov sky’s *Synchronisms, no. 6*) [1]. Electronic improvisation has also been explored, in programs such as Omax, in which the computer actually improvises along with a live instrument [2]. This paper focuses primarily on improvisational or aleatoric elements that affect the large-scale form or other temporal elements of a work. The works discussed below all make use of gestural control in some capacity to determine form and flow of time; that is, in each of these works, an performer interacts with a gestural controller to determine either the overall form of the work, or durations of sections within.

## 2. PRECEDENTS

While improvisation and multiple score interpretation can be observed in many periods throughout history, an exhaustive discussion of improvisatory works is beyond the scope of this paper. Therefore, the practical discussion is preceded with a brief description of some of the more relevant recent works (and scholarship) involving improvisation and/or indeterminacy. The music of John Cage presents an excellent starting point for this discussion, as many of his works contained indeterminate procedures within a modern framework. Many of these works resulted in incredibly diverse interpretations, and Philip Thomas noted that, regarding his *Solo for Piano* (1957-58), “No single interpretation is likely to represent all the material available, and the multiple possibilities open to the interpreter within just a single piece make the *Solo for Piano* a work that is uncontainable: it resists definition, and at micro- and macro- levels the score can be only the beginning of a process, a prompt for action rather than a description of sound” [3]. Thomas further commented on several interpretations of this work, describing that the score was not a definitive element in much of Cage’s work, and that the resulting music was substantially diverse from interpretation to interpretation. When electronics become a consideration, the possibilities for improvisation or indeterminacy increase, especially for those works that involve computing. However, most musicians are not trained to improvise with a computer system (and while growing, pedagogy surrounding the performance of works for instrument and live electronics remains limited). Chapman Welch, who composed several works involving improvisation with an electronic entity, remarked that “the addition of an interactive computer system within an improvisation adds a

new layer of musical input and complexity that is foreign to many improvisers” [4]. Welch describes some of the processes he uses to make improvisation with electronics more approachable for musicians through his work on the piece *Moiré*, for clarinet soloist, computer, and ensemble. The result is highly controlled improvisational blocks, as well as a conductor that gives cues. Welch describes further his future work in storing the musicians’ real time choices in a database and using such data to affect the performance (thereby creating more of a two-way improvisation and effect). This type of two-way approach was also implemented in software designed at IRCAM titled Omax. The software, which combines the software Open Music with Max, “learns in real-time typical features of a musician’s style and plays along with him interactively, giving the flavor of a machine co-improvisation” [5]. The creators of Omax describe it as an “improvisation-oriented musician-machine interaction system that learns in real time from human performers” [6]. Omax has been implemented successfully in numerous performances, and indicates that substantial research in the designing of such systems has been completed, and that improvisation with electronics represents a viable performance genre. Both of the systems discussed here represent effective means of working with improvisation and indeterminate elements specifically in working with a computer. They take a different trajectory however; electronics are incorporated as a primary improvisatory driver, but the nature of the improvisation resembles that of the Cage example discussed above, or another Open Form work. However, some elements, such as Welch’s use of conductor and improvisational feedback loop, do find analogues in elements of my works. The other feature of the works described in this paper is the use of gesture and motion capture systems as an integral component of the improvisation.

### 3. EARLY EXPLORATION – *THE BLACK SWAN*

I created a work, performed twice in late 2013 and presented as a poster at the 2014 ICMC-SMC joint conference, entitled *The Black Swan*, for performance over a network [7]. *The Black Swan* consisted of a graphical score, which was to be realized during performance, a score for “movers”, who to some extent generated the score, and a master score, which transferred performance instructions. All of these scores consist of software components created using Max, and made use of Jacktrip to facilitate the audio transferal over the network [8, 9]. One performance node was tracked with a webcam, with captured movements translated into data using computer vision algorithms developed for Max [10, 11]. While the presence of a score indicates that there is some “composed” or “fixed” element of the piece, the actual form is somewhat improvisatory, as the flow of time through the score is generated during performance. This is executed the following way: one node receives software containing the performance score, which the performer (or performers) reads off of the computer screen in real time. This

performance score is intended to be read by trained musicians, but is in open score format, allowing any ensemble to perform the work. Therefore, the score contains graphical elements, but the basic layout of the score is very familiar to anyone that understands western concert music notation: there is a staff and symbols placed upon it are read left to right, with vertical dimension indicating relative pitch, which is dependent upon the performing instrument’s range. Another node contains instructions for a performer to interact with a motion tracking system; this interaction serves to make small modifications in the advancement of the score, as well as, at pivotal moments, major score changes.



Figure 1 The Black Swan score image

This system allows for the kind of structured improvisation that integrates well with electronic interactivity and networked performance, as the latency involved in network systems changes depending on how far away the two performance nodes are, and what type of network connections they have. This makes precise rhythmic interactions over networks potentially difficult and even problematic [12]. *The Black Swan* allows two performance nodes to be geographically separated and still maintain an engaging performance without performance being compromised by the latency of the network. The motion-tracking node receives instructions for performance from a master controller (a role that in the past has been performed by myself). The master sends text information in real time over the network to the performers on the motion tracking side regarding what type of movements to execute, which direction to move, how much of the body to move, etc. This allows for a significant amount of indeterminacy from performance to performance, as the number of performers on the motion tracking side can vary.



Figure 2 The Black Swan camera tracking software

These performers have a visual display, which indicates a number between one and one hundred, and when the

number reaches one hundred, the score advances to the next section. The tracking used for this is very general, with the computer vision algorithms evaluating the overall direction and amount of movement. The real-time text-based performance instruction creates the performance interest, because the master controller can instruct the movers to do anything. The movement also causes adjustment of the observable score image at any given time. Since the musical performers are instructed very specifically to read the score left to right, this changes the structure of the piece as well. Further, at some point in the score, the motion tracking team is actually instructed to perform some musical sound, which is sent to the score realization team, unbeknownst to them. This instruction is given by the master, and also adds an element of indeterminacy and surprise to the work.

#### **4. THE WOMAN AND THE LYRE PART I: FAYUM FRAGMENTS**

*Fayum Fragments* is an individual poly-work that serves as part of a larger poly-work, entitled *The Woman and the Lyre*, for mezzo-soprano, flute, cello, piano, and live electronics. In this paper, the work *Fayum Fragments* will be discussed as its own individual entity. Specific attention will be paid to the form of *Fayum Fragments*, and how the use of gestural controller is integral to the realization of the overall form of the work. The text that serves as inspiration and setting of the work consists of a series of short Greek fragments extracted from a surviving poem by the Archaic Greek woman Sappho, titled the *Fayum Fragments* [13]. The poem, the second of two appearing in the *Fayum Fragments*, consists of ten lines, of which eight have discernible text. Ten musical miniatures were composed as a response to this poem; each of the miniatures represents the only discernible word or phrase in its line. This actually provided a large amount of inspiration regarding the form of the piece; the meaning behind the poem cannot be determined without the context of the words, and this work seeks to create its own meaning every time it is performed by framing the “words” in a different context. While each of the short instrumental movements can function as a standalone musical work, they are designed for performance either simultaneously or successively (or a mixture of both) as part of larger sections or movements, of which there are three.

##### **4.1 *Fayum Fragments* – form and text**

The overall performance structure of the *Fayum Fragments* is as follows:

- 1) Fragments 1 – Performance of two or three fragments, concluded by a short tutti fragment
- 2) Fragments 2 – Performance of two or three fragments, concluded by a short tutti fragment
- 3) Fragments 3 – Performance of two or three fragments, concluded with silence/end

As discussed previously, each singular fragment is inspired by the sound of a small fragment in Greek derived from a larger, and mostly unreadable, ancient poem. The phrases have been translated into English by Henry T. Wharton to mean: 1) *soul*, 2) *altogether*, 3) *I should be able*, 4) *as long as indeed is to me*, 5) *to flash back*, 6) *fair face*, 7) *stained over*, and 8) *friend*. There were two lines in the original poem of which no text was discernible, and the two tutti (instrumental only) movements in the work represent these by incorporating the (interpreted) rhythmic structure of the lines.

##### **4.2 Aleatoric elements**

*Fayum Fragments* contains ten pre-composed, short musical movements for live performance. However, the order in which they are performed, and the time passing between onsets of each fragment, is determined entirely by the vocalist during the live performance. The vocalist uses a Leap Motion controller, reading off of a gestural score, which instructs her to execute specific hand motions. *Event detection* is used to trigger the onset of new movements once the vocalist performs pre-determined actions. Each time the vocalist triggers a new fragment, the system “looks” for a different circumstance. For example, at one point the software will allow a fragment to be triggered as soon as the vocalist places both hands within the tracking area.

Once this fragment is triggered, however, the software will move on to detect another gestural event. An algorithm is used to determine which movement begins when a new one is triggered. In the very beginning of the *Fayum Fragments*, any of the fragments may be triggered. However, once the work begins, each section only allows for fragments that don’t contain instruments that have already performed or are currently performing to be triggered for performance. Since each fragment consists of a different subset of the instrumental group (or a solo instrument), this allows for the fragments to be ordered somewhat randomly.

##### **4.3 The use of the Leap Motion to generate form**

The vocalist uses a Leap Motion controller to trigger the onset of fragments, and thus determine the form of *Fayum Fragments* during performance. The Leap Motion interfaces with Max using the MRLeap object [14], and tracks the singer’s hand movements along X, Y, and Z-axes, as well as velocity measurements. A gestural score is given to display which gestures the vocalist is to make throughout the performance of each fragment (or fragments). This score is very detailed regarding the type of gestures, their size, which hand to use, and the location on a given axis to start or finish them. However, the vocalist is instructed to perform each section at any speed desired, and is given instructions to leave as much (or as little) time as desired between the onset of fragments. As the order of the fragments is somewhat random, this enables the vocalist to make a decision during the perfor-

mance regarding how quickly to progress through each section. Therefore, the duration of *Fayum Fragments* is highly variable, and can last anywhere from about ten minutes in duration to about twenty minutes in duration.

The work uses a *score-oriented event detection* procedure, since the Max software listens for specific gestures to occur during given points to trigger the performance of each new fragment [15]. Originally, I had considered using specific gestures to control the advancement, or triggering of fragments.

This procedure would have been incredibly simple, as the Leap Motion already has traceable built in gestures, which could have served as triggering elements. However, this would have removed a dimension from the composition that is extremely important, which is that of the singer's role of conductor, or form-determinant, of the work. Since the work contains many other dramatic elements, this was particularly important, and incorporating the flexible and notated gestures allows for the interaction between the singer and the Leap Motion to contain its own drama and narrative that threads through the entirety of the work. This in effect gives a distinct and meaningful form to a work that has so many elements of indeterminacy.

## 5. OPEN SPACE: EXPANDING THE BLACK SWAN

A follow-up work to *The Black Swan* was created in 2015-2016, entitled *Open Space*. This work is intended for performance between two geographically displaced nodes, although it can also be executed between two parties onstage. The work uses motion tracking as a means for score advancement, with detailed instructions a performer must follow, but a loose temporal structure. *Open Space*, however, contains more direct interaction between the two nodes, involving motion tracking in both locations that results in sound and formal advancement.

### 5.1 Space modification node

The space modification node consists of a motion tracking system, which the performer is to engage with. The performer is given loose instructions to explore the physical space of the room or performance area. This exploration serves two purposes: 1) it selects the sound files that will play back and be sent to the sound modification node to perform with, and 2) to spatialize the audio generated by the sound-generating performer. The sound file selection is relatively simple: the video feed is divided into four separate tracking areas. Each of these areas corresponds to one of the environmental soundfiles, and whichever area has the most perceptible motion at any given moment determines which sound file is played back. A crossfade is implemented as well, so that when the performer moves from location to location, the change between sound files is smooth. The location of the performer also results in minor modifications of the environmental sounds, such as delay times (which get longer

as the performer moves further from the centre points), delay feedback (which increases as the performer moves closer to the centre points) and other modifications which simulate distance in some way, such as low pass filter parameters, and output volume. At some point in the piece, the performer will “discover” a location with a bell sound, a discovery, which serves to determine the duration of the piece. Once the bell sound is triggered, the sound modifier is given a very fixed performance score to execute, with instructions that the piece will end at its' completion. The space modifier is also instructed to remain in the space he/she is standing and to only use arm movement to diffuse sound at this time.

### 5.2 Sound modification node

The sound modification in *Open Space* occurs as a result of a performer at a second node also interacting with a camera. The sound-modifying performer is given instructions within the performance materials regarding which gestures, actions, and movements to perform, and a graphical score that they are to execute during a certain moment in the performance. The characteristic of the sound that is generated depends on where, how fast, and how large of a motion a performer makes. However, the exact sounds that are triggered are largely depended on the *space modifier*. Unlike *The Black Swan*, in which the score generator actually advanced the score and determined the form temporally, the space modifier in *Open Space* interacts with a camera to provide the sound generator with an environmental sound file. All of the sounds are derived from recordings taken from various locations, both indoor and outdoor. The sound modifier “explores” each space, as the sound material from the score generator is recorded to a storage buffer that is sent to the sound modifier over the network. The sound modifier then uses gesture and movement to “play” this stored sound as an instrument, using a granular synthesis, and adjusting various parameters of the grains. The computer vision used in *Open Space*, therefore, while making use of the same software as *The Black Swan*, uses the data in much more specific ways, with direct links between movement and resultant sound.

### 5.3 Bell sequence

At an unspecified point during *Open Space*, the space generator will choose to “discover” a space no longer occupied by a location, but by bell sounds. This action causes the work to end; the space modifier can no longer change the environmental sound file. Once the bell sequence is triggered, the space generator serves as sound diffuser, using the motion tracking to move the sounds around the space. The sound generator follows a gestural score during this sequence, and upon completion, the piece is finished.

*Open Space* is therefore, while being quite different in form from *The Black Swan* and *Fayum Fragments*, quite similar in that the overall form and temporal pacing is moderately indeterminate. Although some ele-

ments of *Open Space*, such as the fixed score at the end, are very strictly composed, the overall form of the piece is not. Again, this type of performance is well suited to networked performance environments, in which there may be lagging of both audio and video environments.

## 6. FUTURE DIRECTIONS: SAPPHIC CYCLE

Thus far I have explored those elements of electronic music that pertain mostly to the electronics, and specifically, to gestural control. In all the previously discussed works, gesture control to some effect determines the overall form and duration of the piece. I will now explore some of the compositional (and as an extension, notational) elements I am currently developing that are effective at handling some of the issues that arise when dealing with live electronics in combination with acoustic instruments that have a somewhat fixed score, such as unknown length of effect, the desire for interactivity, and others.

### 6.1 Previous explorations – Redshift

My first experience with a type of aleatoric notation began when I was commissioned to contribute a composition for a spatialized ensemble [16]. This ensemble, which titled itself a *Vertical Orchestra*, performed in various venues, and each time they performed they were spread out throughout the space, rather than close together as a traditional ensemble. This presented some challenges for composition, most notably the lack of visibility of a conductor and other ensemble members. This made the execution of precise, synchronized rhythms difficult, and favoured more flexible notation. To adapt to this medium, I developed a score that was primarily structured improvisation, using boxes to determine the parameters to be performed, and lines and durations to determine how frequently to perform the actions, and whether sound should be continuous or have silence.

The image shows a musical score for four instruments: Fl 1, Fl 2, Sax 1, and Sax 2. The score is written in a standard musical notation style with treble clefs. There are several performance instructions and dynamic markings. For example, 'Occasional dynamic swells; written dynamic should be average' is written above the Fl 1 and Fl 2 staves. 'Timbral trill' is written above the Sax 1 and Sax 2 staves. Dynamic markings include 'p', 'mp', and 'f'. A time signature of ':50' is located at the top right of the score.

Figure 3 Score image from Redshift composition

Many of the issues involved here are also present in networked music (such as lag time, and lack of complete visibility), which contributed to the reason behind my desire to use a more open, aleatoric-inspired notational system for my networked pieces. However, this type of notational thinking also lends itself well to performances in which live instruments interact or perform alongside live electronic systems. It was for this reason that I decided to develop this type of notation further in *Sapphic Cycle*.

### 6.2 Sapphic Cycle – notational elements

*Sapphic Cycle* represents the other part of the larger poly-work *The Woman and The Lyre*. The cycle consists of four songs, whose texts are inspired by fragments of Sappho's texts. These texts are all in English, written in the early twentieth century by Bliss Carman [17]. *Sapphic Cycle* also incorporates motion tracking of the instruments, however, rather than using the Leap Motion, the instrumentalists and the vocalist are tracked via cameras. Creating a flexible notation and interaction system for these songs presented a challenge; I wanted to retain some sense of rhythm, and precise pitch, while also allowing the performers to interact with the electronics in a meaningful way. Therefore, the majority of the cycle is composed in a way that is very close to standard western notation. To allow for flexibility and interactivity with the electronics, I have chosen two means: 1) larger areas with precise rhythmic motives but indeterminate "pausing" time, to allow for the music to breath and the electronics to respond, and 2) areas or blocks in which the performers can "break out" of the metered performance and improvise within the notated parameters. Both of these types of notation allow primarily for temporal flexibility; the pitches, harmonic structures, and rhythmic gestures remain intact. I also have created a number of symbols, which indicate more general parameters; this was done to retain the concept of gesture that permeates the work, and also to provide visuals that are not obtrusive but can convey a multi-dimensional meaning. An example of this is a symbol that is used to indicate that the performers, when within the box, are to perform the notes from any starting point and going in any direction. Words are also used symbolically, to project intentions, emotions, or overall musical concepts. These words can be expressive (such as "lyrically", "harshly", "agitated", etc.) or more action-oriented (such as "as fast as possible"). To separate these intention words from general musical terms used in the score, these words contain a boxed frame.

### 6.3 Sapphic Cycle –implications and interactivity

The two networked pieces described above (*The Black Swan* and *Open Space*) both provide an interesting performance environment for networked concerts and allow the performers to have a large amount of control over the resulting work. However, this type of temporal indeterminacy is not suited for works that require finer control over elements such as pitch, dramatic action, and precise counterpoint. *Fayum Fragments* allows greater control on part of the composer as the scores themselves are fixed, but since the arrangements are indeterminate, this work does not achieve the type of precision a composer may desire in a traditional chamber work. The flexibility in *Sapphic Cycle* is not as dramatic, but at the same time allows the performers to interact with the electronics in a way that they feel is meaningful. The work does not have a diverging or reversible timeline in the way that the previous three works do, yet moving forward, time has some flexibility. As so much of the electronic processing, including spatialization and delay times/feedback, is deter-

mined by the performers' specific gestural actions during performance due to the motion tracking, instructing the performers as to when and how they can modify their own playing gives them greater control over their response and interaction with the electronic sound.

## 7. CONCLUSIONS

Music involving live electronics lends itself to improvisation because of the very changeable nature of electronics, and the ease of implementing indeterminate processes. Several systems have been developed that allow either an acoustic instrumentalist to improvise over electronics, or electronics to improvise in some way with an acoustic instrument. The systems described in this paper serve a very different function than those developments, as the improvisatory and indeterminate elements are very linked to the overall form of works, and especially to the drama and program. These systems integrate gesture for various reasons; such as creating an organic link between musical gesture and electronics (as both are technically effected by the same physical movement at times, such as in *Sapphic Cycle*), or acting as a type of conductor, such as in *Fayum Fragments*. Conducting has its own visual gestural component to it, which can be highly dramatic – *Fayum Fragments* aims to exploit this component. Finally, in networked systems (*The Black Swan* and *Open Space*) gesture is desirable because it allows both an audible and visual link between nodes, thus enhancing the feedback between groups that are not physically close, and increases the potential performer pool. Highly skilled musicians as well as non-musicians explorers in a setting such as an installation can perform both of these networked pieces.

## 8. REFERENCES

- [1] M. Davidovsky, *Synchronisms no. 6*. E.B. Marks Music Corp., 1972.
- [2] G. Assayag, G. Bloch, A. Cont, and S. Dubnov, "Interaction with Machine Improvisation," in *The Structure of Style*, Berlin-Heidelberg, Germany: Springer, 2010.
- [3] P. Thomas, "Understanding Indeterminate Music through Performance: Cage's Solo for Piano," in *Twentieth-Century Music*, vol. 10.01, pp. 91-113, 2013.
- [4] C. Welsh, "Programming Machines and People: Techniques for live improvisation with electronics," in *Leonardo Music Journal*, vol. 20, pp. 25-28, 2010.
- [5] "The Omax Project Page." <<http://omax.ircam.fr/>>
- [6] G. Assayag, et al. "Omax brothers: a dynamic topology of agents for improvisation learning," in *Proceedings of the 1st ACM workshop on Audio and music computing multimedia*. ACM, 2006.
- [7] A. Aska, "The Black Swan: Probable and Improbably Communication Over Local and Geographically Displaced Networked Connections as a Musical Performance System," in *Proceedings of the International Computer Music Conference-Sound and Music Computing Conference*, Athens, Greece, 2014, pp. 553-556.
- [8] Zicarelli, D. Max/MSP Software. San Francisco: Cycling '74. 1997.
- [9] J. Cáceres and C. Chafe, "JackTrip: Under the hood of an engine for network audio," in *Journal of New Music Research*, vol 39.3, pp. 183-187, 2010.
- [10] Pelletier, J. "cv.jit – "Computer Vision for Jitter." <<http://jmpelletier.com/cvjit/>>
- [11] February 15, 2012. Ritter, M., "MR.jit.toolbox", <<http://www.martin-ritter.com/software/maxmsp/mr-jit-toolbox/>>
- [12] Chafe, Chris, et al. "Effect of time delay on ensemble accuracy." *Proceedings of the International Symposium on Musical Acoustics*. Vol. 31. 2004.
- [13] H. T. Wharton, *Sappho: Memior, Text, Selected Renderings, and a Literal Translation*. Amsterdam, Netherlands: Liberac, 1974.
- [14] M. Ritter and A. Aska, "Leap Motion As Expressive Gestural Interface," in *Proceedings of the International Computer Music Conference-Sound and Music Computing Conference*, Athens, Greece, 2014, pp. 659-662.
- [15] S. Emmerson, *Living Electronic Music*. Burlington, VT: Ashgate Publishing Ltd., 2007.
- [16] <http://www.timescolonist.com/entertainment/music/audience-at-centre-of-vertical-orchestra-1.572424>
- [17] B. Carman, *Sappho: One Hundred Lyrics*. A. Moring, The De La More Press, 1906.